X1

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/843,212	04/24/2001	Eric I-Chao Chang	MS1-666US	8802

22801 7590 08/25/2004

LEE & HAYES PLLC
421 W RIVERSIDE AVENUE SUITE 500
SPOKANE, WA 99201

EXAMINER	
BRANT, DMITRY	

ART UNIT	PAPER NUMBER
2655	6

DATE MAILED: 08/25/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)	
	09/843,212	CHANG ET AL.	
	Examiner	Art Unit	
	Dmitry Brant	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *10 June 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-27* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-27* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments with respect to claim 1-27 have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1-5, 10-14, 19-23, are rejected under 35 U.S.C. 103(a) as being unpatentable over Swaminathan (5,704,000).

As per claims 1, 10, 19, Swaminathan et al. discloses a 2-stage pitch set selection:

I. *identifying an initial set of pitch value candidates* within each time segment (frame)... utilizing a first pitch estimation algorithm (10, FIG. 2, FIG. 3, Col. 4, lines 1-64). Here, Swaminathan uses autocorrelation to produce the initial set of pitch candidates P. P consists of elements P(i, j) where i represents the index within P and j represents time instant (Col. 4, line 57). Therefore, for each time instant j (frame j), there is at least i pitch estimates stored in P. Note that several time intervals comprise an

overall signal segment (FIG. 4) so Examiner interpreted *signal segment* of Swaminathan to contain a number of frames, as claimed by this Application.

II. *reducing* the initial set of pitch value candidates to *a select set of pitch value candidates* ... based on re-scoring utilizing a second pitch estimation algorithm (20, FIG. 2, FIG. 5, Col. 6, lines 20-25). Swaminathan uses cost analysis (re-scoring) as the 2nd estimation algorithm to determine the optimal pitch estimate lopt (j) for a given time instant j [frame j] (Col. 6, lines 29-36). The full description of Swaminathan's 2nd selection algorithm is contained in Col. 5, line 23 - Col. 6, 35. Note that the output of the 2nd algorithm is a reduced set of pitch estimates from the first step- one optimal value for each time instant Pj, wherein each time segment contains multiple time instants, so the overall set of P estimates (P total) contains at least several Pj estimates for the different time instants. (Col. 6, lines 49-50)

Swaminathan does not explicitly teach that the second algorithm is performed substantially real time.

However, Swaminathan suggests that the pitch estimation algorithm is intended to cure the problems of modern telephone systems or CELP coders (Col. 1, lines 35-46 and Title). As it is well-known in the art, the coding operations (requiring pitch estimation) in telephone systems is *substantially* real-time, and as a result, Swaminathan's algorithm would have to perform all of its operations very quickly if it was deployed in the telephone system, as suggested by the disclosure.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made that the pitch estimation algorithm of Swaminathan was to

be used in a real-time systems (such as digital telephone systems), so as to improve the modeling and coding of the input speech signal in CELP coders (Col. 1, lines 35-46), and as a result, would have to operate in substantially real-time.

As per claims 2-3, 11-12, 20-21, Swaminathan do not teach calculating transition probability between at least on of the select pitch value candidates of adjacent frames and selecting a pitch value within each frame with the highest transition probability between adjacent frames as the pitch value for the frame.

However, Swaminathan discloses evaluating each of pitch candidates in view of surrounding candidates for other time instants and discarding pitch candidates that are inconsistent with the overall contour of the pitch candidates, while picking the "optimal" candidates that minimize the transitional "cost" (Col. 5, lines 14-22 and Col. 6, lines 20-25). Swaminathan says that pitch estimates which do not correlate to estimates in other time instants (frames) are likely to be caused by noise/errors and should be discarded (Col. 5, lines 19-22). Furthermore, Swaminathan teaches calculating transitional cost for pitch candidates between different time instances (Col. 5, lines 32-47) which measures the distortion between different time instances, and hence, indirectly indicates the transitional probability between such instances (the higher the distortion, the less likely is a valid transition).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Swaminathan to use transitional probabilities instead of cost function (or express transitional probabilities in terms of cost functions,

as it is well-known in the art) to determine whether the pitch estimate is likely to be either valid or invalid, based on the amount of distortion between successive time intervals (frames). The motivation for doing so would have been to cleanse the set of pitch estimates from the estimates that were likely to be caused by noise/errors (Col. 5, lines 14-22 and Col. 6, lines 20-25)

As per claims 4, 13, 22, Swaminathan do not disclose using dynamic programming to calculate a significantly best path between different pitch candidates of adjacent frames.

However, Swaminathan do teach computing optimal path metrics using a formula that determines the optimal path by minimizing the distortion between adjacent pitch estimates (Col. 5, lines 51-63), which would suggest the use of dynamic programming to one skilled in the art. As discussed in rejection for claims 2-3, computing optimal path metrics by minimizing of distortion of inter-frame pitch estimates can be expressed in terms of transitional probabilities by anyone with the ordinary skill in the art. In addition, Examiner takes the official notice that dynamic programming is well-known in the art and is often used as a computationally efficient way of calculating optimal metrics by minimizing path costs, as described in the above formula.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Swaminathan to use dynamic programming to compute the optimal metrics, as it is well-known in the art, in order to utilize a computationally efficient method of computing optimal metrics for Swaminathan's

invention and determining the "optimal" set of pitch estimates (for discussion of how these metrics relate to transitional probabilities, see rejection for claims 2-3).

As per claims 5, 14, 23, Swaminathan teaches smoothing a curve representing the select pitch values over a plurality of frames based on other information (Col.6, lines 40-46, i.e. taking approximate modal average of the optimal pitch candidates, taking into account the possibility that some of these candidates may be in slight error or could suffer from pitch doubling or pitch halving.)

4. Claims 5-6, 14-15, 23-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Swaminathan in view of McCree (6,463,406).

As per claims 5-6, Swaminathan does not disclose pitch smoothing, based on "other information", such as "one or more of an energy value for each frame, a zero crossing rate of the audio content, and/or vocal tract spectrum of the audio content."

McCree teaches smoothing over frequency bands that are chosen from within the vocal track spectrum (Col. 8, lines 58-61 and Col. 7, lines 48-51)

It would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Swaminathan as taught by McCree in order to improve the pitch smoothing process, because smoothing only over the ranges where human speech can occur would de-emphasize the noise coming from other ranges, thus

reducing the possibility that non-speech signals would interfere with the pitch tracking process of human speech signals.

5. Claims 7-8, 16-17, 25-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Swaminathan in view of Cook et al. (5,353,372).

As per claims 7,16,25, Swaminathan does not disclose the use of AMDF for first step of pitch detection and "selecting N near-zero minima pitch values in the audio content as the initial set of pitch value candidates." However, Swaminathan does teach using autocorrelation for first stage pitch candidate selection (14, FIG. 3)

Parsons discloses that AMDF is one of the many different modifications/substitutes of the autocorrelation algorithm used for pitch estimation (see T. Parsons, "Voice and Speech Processing", pages 202-203). Parsons teach that AMDF pitch detection necessarily involves estimating period using the location of the minimal pitch values, as near-nulls occur at or around integer multiples of the period. (see Figure 8-5).

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Swaminathan as taught by Parsons to use AMDF instead of standard autocorrelation for the first stage of pitch estimation, in order to improve the effectiveness of the pitch estimation algorithm, because AMDF is well-known alternative to the standard autocorrelation algorithm and is also well-known to be

used for speech coding applications, such as telephony (Parsons, page 203, first paragraph)

As per claims 8, 17, 26, neither Swaminathan nor Parsons teach setting N to 288.

It would have been obvious to one of ordinary skill in the art at the time the invention was made that AMDF detector would require a sufficient number of zero samples in order to produce a reasonable approximation of pitch. This happens because tested signal is often not truly periodic and pitch nulls exist between integer values of the period. As a result, selecting a larger set of test values, such as 288, would improve the reliability of pitch estimation.

6. Claims 9, 18, 27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Swaminathan .

Swaminathan does not disclose using NCCF for the second step of the pitch estimation process, where the originally selected set of pitch values is further reduced with NCCF to a "select set of pitch values."

However, it would have been obvious to one of ordinary skill in the art at the time the invention was made that NCCF is a commonly used method of computing correlation within a group of signals. Therefore, one could further limit the first set of pitch estimates by first computing NCCF for each signal in the set and then choosing

the signals that have the highest NCC values, because these signals are more likely to estimate the correct value of pitch than signals that do not correlate with the group.

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Swaminathan to use NCCF (instead of inter-frame cost computation) on the first set of estimated pitch values and pick M best cross-correlated values as possible pitch estimates, because using highly cross-correlated estimates will further improve the probability that the picked estimates correspond to the actual pitch value.

Allowable Subject Matter

7. The following claim 1 is drafted by the examiner (including limitations of claims 7 and 9) considered to distinguish patentably over the art of record in this application is presented to applicant for consideration:

1. A method comprising: identifying an initial set of pitch value candidates within each frame of a plurality of frames of received audio content utilizing a first pitch estimation algorithm; and reducing the initial set of pitch value candidates to a select set of pitch value candidates based, at least in part, on pitch value re-scoring utilizing a second pitch estimation algorithm, wherein the select set of pitch values are selected in substantially real-time;

wherein identifying the initial set of pitch value candidates within each frame comprises: passing each frame of audio content through an average magnitude

difference function (AMDF); and selecting N near-zero minima pitch values in the audio content as the initial set of pitch value candidates; and

wherein identifying a select set of pitch values comprises: generating a local score for each of the initial set of pitch value candidates utilizing a normalized cross-correlation function (NCCF); and selecting M pitch value candidates with the highest local score.

Examiner believes that this claim properly covers Applicant's invention, as it was presented in the IDS article "Large Vocabulary Mandarin Speech Recognition with Different Approaches in Modeling Tones."

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Two step pitch estimation

[1] Secrest et al. (4,731,846)

[2] Doddington et al. (4,696,038)

[3] Koniklijke et al. (WO 99/59138)

[4] Crepy et al. (4,924,508)

General pitch estimation

[5] Yeldener (6,456,965)

[6] Redkov et al. (6,496,797)

[7] Thomas Parsons, "Voice and Speech Processing," pp. 199-203, McGraw-Hill (1987)

[8] D. Tuffelli, "A pitch detection algorithm with hypothesis and test strategy by means of fast surface AMDF," Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84. ,Volume: 9 , Mar 1984, Pages: 81 - 84

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dmitry Brant whose telephone number is (703) 305-8954. The examiner can normally be reached on Mon. - Fri. (8:30am - 5pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to Tech Center 2600 receptionist whose telephone number is (703) 305- 4700.

DB

8/18/04

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER